MOLECULAR ECOLOGY

Molecular Ecology (2016) **25**, 3019–3032 doi: 10.1111/mec.13646

Investigating population differentiation in a major African agricultural pest: evidence from geometric morphometrics and connectivity suggests high invasion potential

M. KARSTEN,* P. ADDISON,* B. JANSEN VAN VUUREN† and J. S. TERBLANCHE‡

*Department of Conservation Ecology and Entomology, Stellenbosch University, Private Bag X1, Matieland 7602, South Africa, †Department of Zoology, Molecular Zoology Laboratory, University of Johannesburg, PO Box 524, Auckland Park 2006, South Africa, ‡Department of Conservation Ecology and Entomology, Centre for Invasion Biology, Stellenbosch University, Private Bag X1, Matieland 7602, South Africa

Abstract

The distribution, spatial pattern and population dynamics of a species can be influenced by differences in the environment across its range. Spatial variation in climatic conditions can cause local populations to undergo disruptive selection and ultimately result in local adaptation. However, local adaptation can be constrained by gene flow and may favour resident individuals over migrants—both are factors critical to the assessment of invasion potential. The Natal fruit fly (*Ceratitis rosa*) is a major agricultural pest in Africa with a history of island invasions, although its range is largely restricted to south east Africa. Across Africa, *C. rosa* is genetically structured into two clusters (R1 and R2), with these clusters occurring sympatrically in the north of South Africa. The spatial distribution of these genotypic clusters remains unexamined despite their importance for understanding the pest's invasion potential. Here, *C. rosa*, sampled from 22 South African locations, were genotyped at 11 polymorphic microsatellite loci and assessed morphologically using geometric morphometric wing shape analyses to investigate patterns of population structure and determine connectedness of pest-occupied sites. Our results show little to no intraspecific (population) differentiation, high population connectivity, high effective population sizes and only one morphological type (R2) within South Africa. The absence of the R1 morphotype at sites where it was previously found may be a consequence of differences in thermal niches of the two morphotypes. Overall, our results suggest high invasion potential of this species, that area-wide pest management should be undertaken on a country-wide scale, and that border control is critical to preventing further invasions.

Keywords: dispersal potential, isolation-by-distance, isolation-by-environment, migration, population dynamics

Received 15 September 2014; revision received 7 March 2016; accepted 6 April 2016

Introduction

Environmental variation across the geographic range of a species influences population dynamics (Forester *et al.* 2016) and climatic conditions can result in disruptive selection driving local population adaptation (Richter-Boix *et al.* 2010). Although it is generally accepted that gene flow constrains adaptation (Lenormand 2002), several studies have shown that local adaptation can occur despite relatively high gene flow (e.g. Von Wettberg *et al.* 2008; Richter-Boix *et al.* 2010; Berthouly-Salazar *et al.* 2013). Several aspects of demography influence the likelihood of local adaptation and population- or individual-level genetic variance. Dispersal rates of individuals between populations influence gene flow

Correspondence: John S. Terblanche, Fax: +27 21 808 2546; E-mail: jst@sun.ac.za

directly, however, other factors critical to population differentiation include how successful and fecund migrants are in the novel habitat when mating with resident individuals (Kawecki & Ebert 2004). Local adaptation may therefore provide a selective advantage to residents over migrants. The relative roles of biotic and abiotic factors in contributing to population differentiation and rates of local adaptation are highly contentious, with implications for a wide range of ecological and evolutionary processes, including micro- and macroevolution (e.g. speciation) (see e.g. Berthouly-Salazar *et al.* 2012). Biological invasions provide a unique opportunity to test key theories surrounding factors potentially limiting (or contributing to) population structure or morphological adaptations, and assess rates of local population differentiation. Arguably, the most prevalent and widespread hypothesis is isolation-by-distance (IBD), although recent meta-analysis suggests isolation-by-environment may be equally important when it is considered (Sexton *et al.* 2014). For biological invasions, a deeper understanding of how demographics contribute to (or limit) population establishment and spread is critical. Indeed, both environmental and demographic factors are emphasized in recent theoretical models of population establishment of invasive animals (Berthouly-Salazar *et al.* 2012; Duncan *et al.* 2014). Despite the importance of propagule pressure in the initial establishment of a species in a novel environment (Lockwood *et al.* 2009), genetic processes are nevertheless critical to the abundance and spread of populations after establishment (Szűcs *et al.* 2014).

Many fruit flies of the Tephritidae (Diptera) are successful invaders worldwide (White & Elson-Harris 1994; Hill *et al.* 2016). Two of the most economically significant species in Africa are the Mediterranean fruit fly [*Ceratitis capitata* (Weidemann)] and the Natal fruit fly (*Ceratitis rosa* Karsch) (White & Elson-Harris 1994). Unlike *C. capitata* that has become invasive in many locations outside of its native range, *C. rosa* is largely restricted to eastern and southern Africa, but has successfully spread to the Indian Ocean islands of La Réunion and Mauritius and even outcompete *C. capitata* in some instances (White *et al.* 2001; Duyck *et al.* 2006; De Meyer *et al.* 2008). Since the life-history characteristics, polyphagy and thermal niche of the two species are similar (Duyck & Quilici 2002; De Meyer *et al.* 2008), this raises major concerns for the international fruit industry that *C. rosa* may possess high invasion and competitive potential. These concerns are further underscored by *C. rosa*'s wide host range (>100 hosts), similar to that of *C. capitata* (White & Elson-Harris 1994; De Meyer *et al.* 2002; Copeland *et al.* 2006). *Ceratitis rosa* may therefore pose a high risk of establishment outside of its native range (De Meyer *et al.* 2008; De Villiers *et al.* 2013a).

Molecular information provides indirect estimates of dispersal (gene flow) between populations. The advantage of such an approach is that estimates are averaged over several generations, overcoming some of the drawbacks of using more direct measurements (e.g. mark-and-recapture methods) (Bohonak 1999). Comparisons between direct measures of migration and indirect measures often provide different estimates, with gene flow estimates based on DNA data typically being higher (Bohonak 1999; but see Bouyer *et al.* 2009). This discrepancy is clear in *C. capitata* where direct measures of life-time migration distances are, on average, between 0.5 and 9.5 km (Meats & Smallridge 2007) and reported estimates based on indirect measures are in the order of hundreds of kilometres (e.g. Karsten *et al.* 2013). Estimates of dispersal are not available for *C. rosa*, but based on the generally leptokurtic dispersal pattern of tephritids and the overall morphological similarity between *Ceratitis* species, the distances for *C. rosa* are probably in the same range as that of *C. capitata*. Geometric morphometric measurements have been shown to be effective in discriminating significant ecotypes or groups at intra and interspecific levels (Gilchrist & Crisafulli 2006; Bouyer *et al.* 2007; Marsteller *et al.* 2009; Schutze *et al.* 2012a,b). Geometric morphometrics quantifies shape variation between homologous structures (e.g. key landmarks on wings) and then compares this shape data between groups or individuals (Rohlf & Marcus 1993; Rohlf 1999). Furthermore, morphological data may be amongst the first characters to show differences between populations (Bouyer *et al.* 2007) where population genetic structure estimates may be unable to resolve early differentiation among populations.

Microsatellite markers are routinely used to document genetic structure in tephritids (e.g. Bonizzoni *et al.* 2001, 2004; Baliraine *et al.* 2004; Nardi *et al.* 2005; Aketarawong *et al.* 2007, 2014; Gilchrist & Meats 2009; Virgilio *et al.* 2010; Karsten *et al.* 2015). Notwithstanding the successful implementation of molecular markers and despite the potential importance of *C. rosa* as a major agricultural pest including growing concerns for further spread and establishment within and outside of Africa, the population genetic differentiation of *C. rosa* has been surprisingly poorly studied. Early work reported some genetic differentiation between Africa and island populations of Mauritius and La Réunion (Baliraine *et al.* 2004). Studies reported for *C. rosa* to date have focused mainly on molecular diagnostics (Douglas & Haymer 2001; Barr *et al.* 2006), taxonomic status and phylogenetic relationships with other Tephritidae (Barr & McPheron 2006; Virgilio *et al.* 2008). *Ceratitis rosa* forms part of a complex of three closely-related species called the *Ceratitis* 'FAR' complex (Diptera: Tephritidae) which also includes *Ceratitis*

fasciventris (Bezzi) and *Ceratitis anonae* Graham (Virgilio *et al.* 2008). Based on sixteen microsatellites (FAR 1-FAR16; Delatte *et al.* 2013), five genotypic clusters were characterized in these three species with intercluster divergences being equal to, or higher than, divergences between recognized species (Virgilio *et al.* 2013). Within *C. rosa*, two genotypic clusters were identified as R1 and R2 with largely overlapping geographic ranges (Virgilio *et al.* 2013); little to no genetic differentiation exists within these two genetic clusters. The genetic differences between the clusters are partly supported by morphological differences in the shape and feathering of the mid tibia of the males (De Meyer *et al.* 2015) such that they may be considered separate morphotypes, but with no definite distinguishing characters identified for the females as yet (Virgilio *et al.* 2013). Separation between the two morphotypes is further supported by differences in their cuticular hydrocarbon profiles (Vaníčková *et al.* 2015). A recent study comparing their thermal development requirements suggested that the R1 morphotype was better adapted to warmer temperatures than the R2 morphotype (Tanga *et al.* 2015). These results are underscored by the geographic distribution of the two morphotypes in Tanzania along an altitudinal transect (Mwatawala *et al.* 2015). Although both morphotypes are found throughout the region, the R1 morphotype is more abundant at lower altitudes compared to the R2 morphotype, the latter of which is more abundant at higher altitudes.

Understanding of the distribution of the two morphotypes within South Africa, as well as their regional scale population structure is critical as this information provides essential knowledge for Integrated Pest Management (IPM) strategies and management of biological invasions within and between countries. In South Africa, a unique situation exists as the two morphotypes (R1 and R2) occur sympatrically in the northern regions of the country (see fig. 3 in Virgilio *et al.* 2013). Moreover, morphological inspection of the legs of *C. rosa* males from South Africa indicated additional locations at which both morphotypes (albeit R1 at low abundance) occur sympatrically in the north of South Africa, especially the Mpumalanga and Limpopo Provinces (Tzaneen, Rustenburg, Komatiepoort, Nelspruit, Marblehall) (M. de Meyer, personal communication). However, given the small number of locations included for South African *C. rosa* in the Virgilio *et al.* (2013) study, the extent of overlap, as well as the precise geographic distribution of each cluster, remains largely unknown. To address these important knowledge gaps, in this study, we generated both molecular and morphological data for the same individuals, resulting in two parallel but independent data sets. Surprisingly, few studies have adopted this approach despite the clear potential value thereof (but see Bouyer *et al.* 2007; Francoy *et al.* 2009; Bomfim *et al.* 2011; Krosch *et al.* 2013; Francuski *et al.* 2014; Schutze *et al.* 2015a,b).

The broad aims of our study are therefore to: (i) determine the distribution of the *C. rosa* R1 and R2 genotypic clusters across South Africa with possible regions of overlap, (ii) determine the phylogeographic patterns for these two genotypic clusters including possible evolutionary breaks, and finally, (iii) determine the levels of connectivity (gene flow) between populations or sampling localities. We hypothesized that the majority of localities across South Africa will have only one genotypic cluster, but based on the earlier work by Virgilio *et al.* (2013), that there may be areas of overlap for the two clusters in the northern regions of the country. Here, we make use of both morphological and molecular markers to identify the pattern of population structuring of the two morphotypes within South Africa and to investigate the connectedness of pest-occupied sites, with the ultimate aim of informing area-wide pest management recommendations and a better understanding of the species' invasion potential across Africa.

Materials and methods

Sampling sites and fly collection

Ceratitis rosa individuals were collected from 22 sites throughout South Africa (Fig. 1; Table 1) using Bucket traps (Chempac, Paarl, South Africa) baited with the three-component attractant, Biolure® (Chempac, Paarl, South Africa). Flies from different traps were handled separately and were sexed and identified with a standard tephritid key using a stereomicroscope. After identification, both wings (left and right) were removed for morphological analyses and DNA was extracted from the bodies using a DNeasy® tissue kit (QIAGEN Inc.). Following DNA extractions, exoskeletons were washed and stored in absolute ethanol for morphotype identification.

Microsatellite genotyping and analysis

All individuals collected ($N = 458$) were genotyped for 12 microsatellite markers (FAR1–4, 6, 8, 9, 11, 12, 14–16; Delatte *et al.* 2013), although only 11 markers were included in analyses for reasons outlined below. A PCR reaction consisting of 2 μL of DNA (~30 ng), 2 μL of primer mix (2 mM), 6 μL of 2× QIAGEN Multiplex Master Mix and 1 μL water (Multiplex PCR kit; QIAGEN Inc.) were set up for each individual. PCR conditions followed Delatte *et al.* (2013). For each plate, a positive control was included to check for amplification consistency when reading plates. Samples were

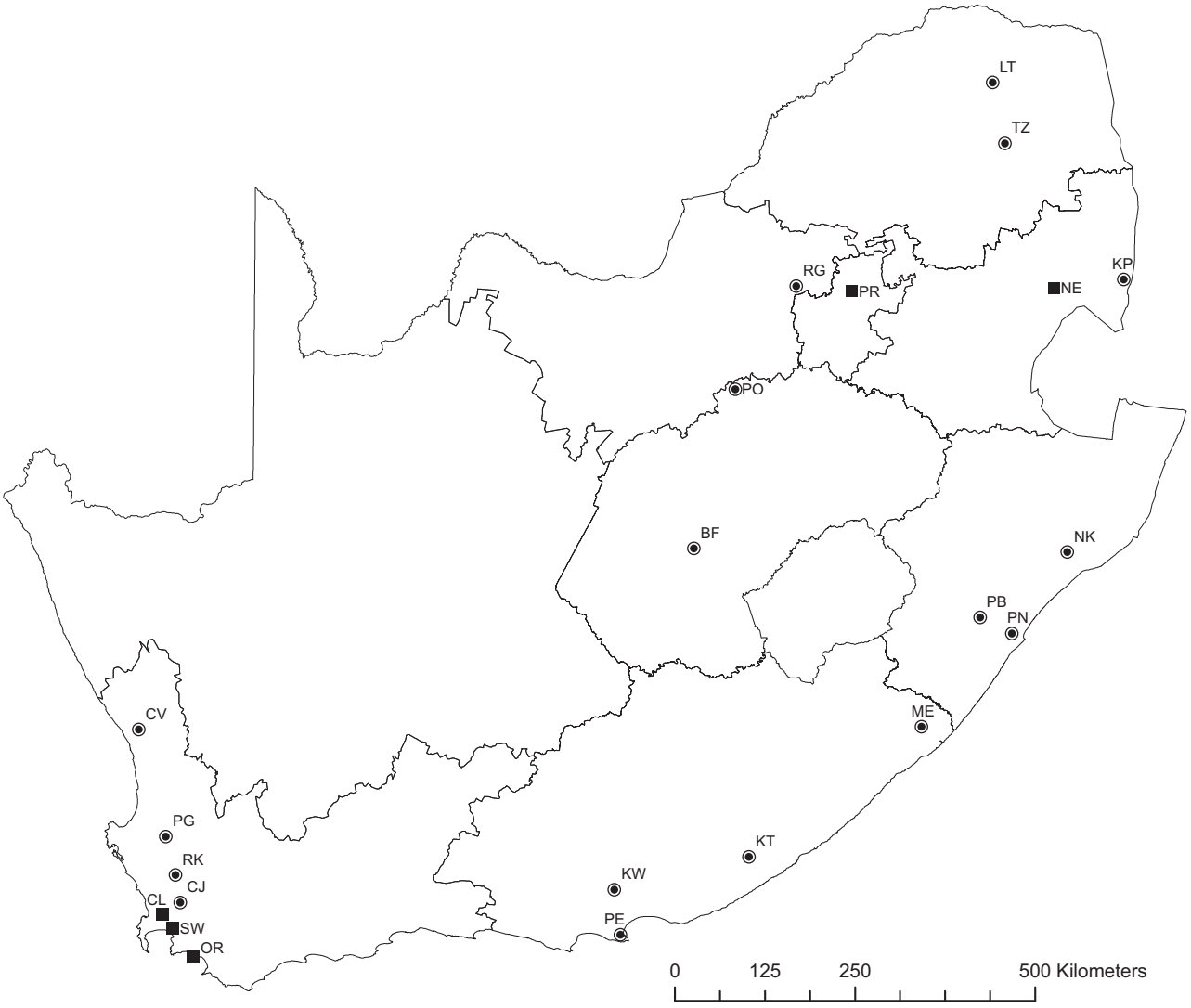

Fig. 1 The 22 sampling sites for *Ceratitis rosa* in South Africa. The five square locations correspond to overlapping sampling localities with Virgilio *et al.* (2013). Population locality names are given in full in Table 1.

genotyped on an ABI 3130 Automated Sequencer (Applied Biosystems, Foster City, CA, USA) and the alleles scored using GENEMAPPER v3.7 software (Applied Biosystems).

Tests for linkage disequilibrium (LD) and departures from Hardy–Weinberg equilibrium (HWE) were conducted in GENEPOP v4.0.10 and statistical significance determined through 1000 permutations (Raymond & Rousset 1995; Rousset 2008). Levels of significance were adjusted using False Discovery Rates (*Q*-VALUE; Storey 2002). LD was observed for one of the markers (FAR9 with FAR3 and FAR15) and FAR9 was therefore excluded from subsequent analyses. To test for the presence of null alleles (A_n), their frequencies in each of the populations for each of the loci were estimated in FREENA v1.0 (Chapuis & Estoup 2007) using the algorithm from Dempster *et al.* (1977). For each locus and for each site, we calculated expected heterozygosity (H_E, the expected allele frequencies under HWE), observed heterozygosity (H_O, the actual heterozygosity in a population measured), the number of alleles (N_A) and the inbreeding coefficient (F_{IS}, the proportion of variance from a population that is contained in an individual) (GENETIX v4.05.2; Belkhir *et al.* 1996–2004; GENALEX v6.5; Peakall & Smouse 2006, 2012).

We explored the demographic history of populations sampled using the MIGRAINE software (http://kimura.univ-montp2.fr/~rousset/Migraine.htm) implementing the 'single population with variable population sizes' model (Leblois *et al.* 2014). Due to a lack of structuring in the data (see Results), all populations were pooled and runs consisted of 2000–10 000 trees, 2400–5000

Table 1 The locations of *Ceratitis rosa* sampling in South Africa with date of collection, type of site where collected (O = orchard, HG = home garden), sample size (*N*), number of alleles (N_A), expected (H_E) and observed (H_O) heterozygosity (±standard error), the inbreeding coefficient (F_{IS}), mean null allele frequency (A_n; Dempster *et al.* 1977; *SD* in parentheses) and effective population size (N_e; lower and upper confidence interval values shown in parentheses) determined from 11 microsatellite markers

Location	ID	Date	Site	GPS coordinates		*N* (genetics/ wings)	N_A	H_E	H_O	F_{IS}	A_n	N_e
				Latitude	Longitude							
Bloemfontein	BF	2007	HG	−29.09	26.28	27/12	6.55	0.636 ± 0.168	0.548 ± 0.176	0.16	0.064 (0.088)	137.7 (31.4–∞)
King Williams Town	KT	2006	HG	−32.87	27.391	14/11	5.73	0.611 ± 0.174	0.509 ± 0.237	0.21	0.080 (0.098)	45.1 (11–∞)
Kirkwood	KW	2006	HG	−33.4	25.44	28/13	6.27	0.607 ± 0.189	0.520 ± 0.235	0.16	0.071 (0.116)	771.6 (44–∞)
Komatiepoort	KP	2007	HG	−25.35	31.914	28/12	7.36	0.654 ± 0.133	0.567 ± 0.169	0.16	0.070 (0.082)	∞ (269.7–∞)
Louis Trichardt	LT	2011	O	−23.07	29.93	16/12	5.36	0.642 ± 0.119	0.570 ± 0.235	0.15	0.061 (0.111)	119.5 (25.1–∞)
Lutzville	CV	2010	HG	−31.56	18.359	10/6	4.36	0.597 ± 0.171	0.510 ± 0.256	0.2	0.080 (0.127)	∞ (16.3–∞)
Mt. Edgecombe	ME	2010	HG	−31.08	29.722	5/5	3.82	0.607 ± 0.132	0.582 ± 0.244	0.15	0.066 (0.085)	∞ (32.5–∞)
Nelspruit	NE	2010	O	−25.53	30.991	30/21	6.91	0.660 ± 0.096	0.591 ± 0.172	0.12	0.062 (0.075)	651 (62.2–∞)
Nkwalini	NK	2007	O	−28.75	31.524	8/10	4.09	0.577 ± 0.193	0.519 ± 0.270	0.18	0.071 (0.121)	47 (7.6–∞)
Onrus River	OR	2007	HG	−34.41	19.172	23/12	5.82	0.645 ± 0.125	0.547 ± 0.165	0.18	0.061 (0.082)	∞ (47.4–∞)
Paarl	CJ	2007	HG	−33.73	18.962	26/15	5.46	0.622 ± 0.131	0.517 ± 0.170	0.19	0.081 (0.094)	∞ (70.7–∞)
Pietermaritzburg	PB	2007	HG	−29.66	30.402	28/15	6.46	0.643 ± 0.110	0.536 ± 0.192	0.19	0.068 (0.088)	∞ (140.4–∞)
Piketberg	PG	2007	HG	−32.9	18.751	23/8	5.36	0.640 ± 0.115	0.510 ± 0.188	0.23	0.084 (0.094)	51.7 (21.7–∞)
Pinetown	PN	2010	O	−29.82	30.871	10/7	4.91	0.689 ± 0.140	0.532 ± 0.188	0.28	0.101 (0.098)	47.2 (10.7–∞)
Port Elizabeth	PE	2010	O	−33.96	25.571	28/16	7.73	0.698 ± 0.115	0.600 ± 0.154	0.16	0.057 (0.098)	∞ (85.6–∞)
Potchefstroom	PO	2010	O	−27.08	26.723	12/10	5.36	0.624 ± 0.153	0.540 ± 0.206	0.18	0.065 (0.108)	57.7 (14.1–∞)
Pretoria	PR	2010	HG	−25.78	28.244	23/17	6.73	0.662 ± 0.148	0.557 ± 0.198	0.18	0.065 (0.105)	∞ (48.0–∞)
Riebeeck Kasteel	RK	2008	HG	−33.39	18.891	18/6	4.55	0.588 ± 0.182	0.416 ± 0.215	0.32	0.111 (0.102)	∞ (31.5–∞)
Rustenburg	RG	2007	HG	−25.76	27.474	27/8	6.91	0.661 ± 0.116	0.544 ± 0.140	0.2	0.077 (0.087)	∞ (87.9–∞)
Somerset West	SW	2007	HG	−34.06	18.862	28/15	5.82	0.643 ± 0.127	0.514 ± 0.219	0.22	0.083 (0.120)	127.6 (37.5–∞)
Stellenbosch	CL	2010	O	−33.88	18.741	25/25	5.91	0.640 ± 0.152	0.448 ± 0.202	0.32	0.118 (0.116)	∞ (41.9–∞)
Tzaneen	TZ	2007	HG	−23.81	30.158	23/10	6.82	0.689 ± 0.140	0.531 ± 0.188	0.16	0.072 (0.070)	82.5 (30.6–∞)

points and four iterations, which produced point estimates and 95% confidence intervals for actual and ancestral θ (scaled population sizes) and *D* (scaled duration). We calculated effective population size (N_e) using the LD method (Waples & Do 2008) as a single-sample estimator in NeESTIMATOR v2 (Do *et al.* 2014). We screened out alleles occurring at low frequencies (P_{crit} = 0.05) and implemented the jackknife method (Waples & Do 2008) to determine confidence intervals for all calculations. The hypothesis of IBD was assessed by the linear correlation of geographic distance (ln distance in km) with genetic distance [$F_{ST}/(1-F_{ST})$] (GENEPOP v4.0.10; Raymond & Rousset 1995; Rousset 2008).

Population differentiation was assessed using the Bayesian clustering method implemented in the program STRUCTURE v2.3.4 (Pritchard *et al.* 2000; Falush *et al.* 2003) as well as estimates of pairwise F_{ST} values in MICROSATELLITE ANALYSER v4.05 (MSA; Dieringer & Schlötterer 2003). Multilocus genotypes were assigned in STRUCTURE without including prior spatial population information. We ran STRUCTURE for 3 000 000 MCMC permutations (where statistical parameters reached stability) with 1 500 000 runs discarded as burn-in from each run for five independent runs for each *K* value, varying respectively between 1 and 22 (the maximum number of sampled sites). We implemented the admixture model with correlated allele frequencies, as we expect common ancestry, and we allowed the Dirichlet distribution of allelic frequencies (λ) to be inferred separately for each site. Choosing the optimal number for *K* is up to the user based on the data. Inspecting the log-probabilities of several values of *K* is thought to be a good estimator of the true value of *K* (Pritchard *et al.* 2000). This method can, however, be problematic when there is more than one high value of Ln *P*(*X*/*K*) and we therefore also inspected the delta *K* value calculated according to Evanno *et al.* (2005) implemented in the online resource STRUCTURE HARVESTER (Earl & Von Holdt 2012). This method also has some drawbacks, more specifically the inability to assess *K* = 1 as the optimal number of clusters. Runs were averaged in CLUMPP v1.1.2 (Jakobsson & Rosenberg 2007) and visualized in DISTRUCT v1.1 (Rosenberg 2004). The hierarchical structuring of populations were investigated using an AMOVA based on R_{ST} estimates in ARLEQUIN v3.5.1.2 (Excoffier & Lischer 2010) using 10 000 permutations.

Wing morphometrics

Wings (left and right) were permanently mounted on microscope slides using Entellan®, a rapid embedding agent. We chose fourteen homologous type I landmarks (Bookstein 1991) (Appendix S1, Supporting information). Twelve of the chosen landmarks were based on Schutze *et al.* (2012a,b); the remaining two were chosen to improve coverage of the wing. Each wing was imaged using a Leica MZ 16A microscope camera. The 14 landmarks were superimposed on the images of the wings and digitized in TPSDIG version 2.17 (Rohlf 2005). To assess possible measurement errors in the data, we investigated digitizing error (positioning of the landmarks) and imaging error (owing to taking the picture). For this purpose, we constructed an additional data set comprising a subset of the individuals from the complete data (henceforth known as the error data set). For each individual, the wings were photographed twice (imaging error) and landmarks were placed on both sets of images of wings independently (digitizing error).

After acquisition of the *x*, *y* coordinates for each set of landmarks, the data were imported into MORPHOJ version 1.05c for shape analysis (Klingenberg 2011). In MORPHOJ, we performed a full Procrustes fit to extract shape information and ran a Generalized Procrustes ANOVA (GPA) on both the error data set and the complete data set. To assess asymmetry in our data set, we included 'side' and 'individuals' as a main effect in the GPA. A covariance matrix for the complete data was used for calculating a Principle Components Analysis (PCA) based on an averaged data set (left and right wing), followed by a Canonical Variate Analysis (CVA). Finally, to test for IBD, we regressed the Mahalanobis distance based on the shape variable (from the CVA) against geographic distance (km).

Results

Microsatellite DNA analyses

Analyses of the 11 polymorphic microsatellite markers in *C. rosa* for 22 geographic localities showed that all of the populations deviated from HWE. Marginal but significant levels of inbreeding (F_{IS} = 0.122–0.323) (Table 1) was evident within sites, indicative of nonrandom mating (similar to results reported by Virgilio *et al.* 2013). The number of alleles (N_A) for the different loci varied, on average, between 3.81 and 7.73 (Table 1). The average null allele frequency per locus ranged between 0.010 (SD = 0.017) at FAR14 and 0.272 (SD = 0.073) at FAR1 (Table 1; Appendix S2, Supporting information). All loci were included in all analyses given that F_{ST} estimates and genetic distances are largely unbiased when population structure is absent (see below) (Chapuis & Estoup 2007). Genetic diversity, as measured by mean expected heterozygosity (H_E), ranged between 0.577 (Nkwalini) and 0.698 (Port Elizabeth) (Table 1).

Results from MIGRAINE for the pooled populations indicated a significant population expansion, with ancestral θ = 2.986 (1.587–4.014), actual θ = 38.81 (8.805–144.100)

and D = 0.000896 (0.000114–0.029400) (Fig. 2A,B). Effective population size estimates (N_e) ranged from 45.1 to infinity with large confidence intervals (Table 1).

Tests for IBD patterns in this data set showed no correlation between genetic distance and geographic distance (r = 0.088, P = 0.184; Appendix S3, Supporting information). Results from the Bayesian clustering of individuals showed a peak in the ΔK at K = 2 (Fig. 3A). Taking into account that the approach from Evanno *et al.* (2005) cannot account for K = 1, we also examined the posterior probabilities (Ln $P(X/K)$) which was highest for K = 1. This result was confirmed by STRUCTURE analyses including only those localities within our data set included in both the present and Virgilio *et al.* (2013) study (Fig. 3B).

Results from AMOVA indicated that variation was overwhelmingly partitioned within localities, with 87.3% of variation accounted for by the within-population component (F_{ST} = 0.127; P < 0.0001; Appendix S4, Supporting information). Only 6.9% of pairwise F_{ST} comparisons between sites were significant (Appendix S5, Supporting information). Most of the significant differences were associated with the Kirkwood site, which also accounted for the significant F_{ST} results.

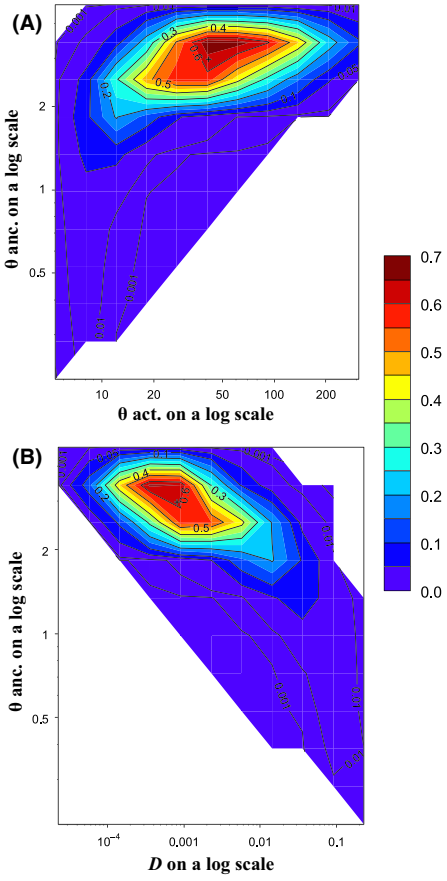

Fig. 2 Results from MIGRAINE showing profile likelihood surfaces for ancestral (anc. θ) and actual (act. θ) effective population sizes as well as timing of the demographic history events (D).

Morphometrics: wing shape

Results from the Procrustes ANOVA (Appendix S4, Supporting information) based on the error data set showed that there was no significant measurement error (both digitizing and imaging error) (F = 1.03, P = 0.302, d.f. = 864). Two hundred and sixty-five individuals were included in the full data set (Table 1) and a PCA for all individuals showed clear distinction between males and females (Fig. 4A, F = 85.88, P < 0.0001). In subsequent analyses, we therefore separated the two sexes. For both males (F = 4.15, P < 0.0001) and females (F = 5.96, P < 0.0001), there was a significant side effect, indicative of a consistent directional asymmetry. Directional asymmetry is found as a result of consistent differences between the two sides (left and right) of objects and individuals and is therefore not a property of the individual but of the population. The CVA plots for both males and females based on the first two canonical variates showed no clear discrimination among 21 locations (in two locations we only obtained single-sex samples; in Potchefstroom, there were no male individuals and in King Williams Town no females) (Fig. 4B,C). No pattern of IBD was detected when Mahalanobis distance was plotted against geographic distance for both males (r = 0.014, P = 0.834) or females (r = 0.016, P = 0.805) (Appendix S6, Supporting information).

Discussion

The population genetics of several insect pest species have been widely studied using molecular tools as well as Bayesian approaches. In contrast, surprisingly few studies have combined genetic data with phenotypic data (e.g. geometric morphometrics) to gain insight into potential evolutionary divergence at the population level. Here, we use this powerful approach to characterize genetic diversity and genetic structure alongside population discrimination based on geometric morphometrics. These results largely matched our hypothesis of a single genotypic cluster (R2) over the majority of the country. Unexpectedly, despite high sampling effort, we were unable to identify any individuals of the second morphotype (R1) at locations where it has previously been reported (see Virgilio *et al.* 2013 and Introduction). Furthermore, we found high levels of genetic diversity, high effective population sizes and unexpected low levels of intraspecific population differentiation between all sampling locations. In the

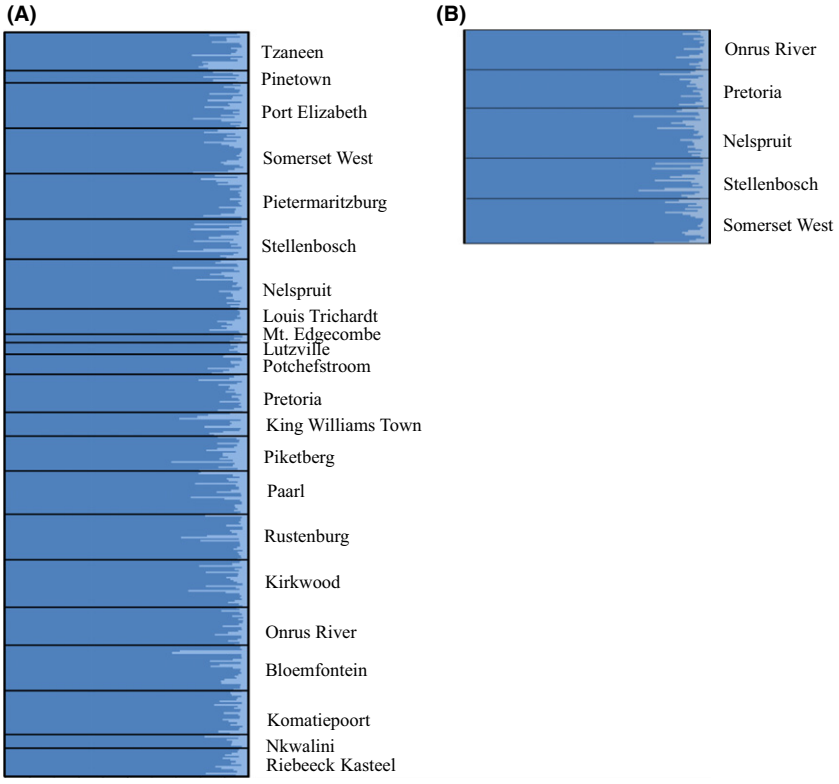

Fig. 3 STRUCTURE results for (A) 458 *Ceratitis rosa* individuals from 22 South African populations for $K = 2$ (B) a subset of locations from this study for 129 *C. rosa* individuals from five sampling locations matching those of Virgilio *et al.* (2013).

complementary morphometric approach, geometric morphometric analyses also failed to show interspecific differentiation between sampling locations suggesting that the lack of population differentiation is a conclusion that is likely robust and not confounded by the choice of molecular approach (e.g. gene flow may be hard to estimate accurately for large effective population sizes; Chapuis *et al.* 2011). These major findings are discussed in detail below.

Overall genetic diversity estimates for *Ceratitis rosa* in South Africa were high, but comparable to those of other studies for South Africa (Baliraine *et al.* 2004; Virgilio *et al.* 2013) and other African populations, for example, Kenya (Baliraine *et al.* 2004; Virgilio *et al.* 2013) and Mozambique (Virgilio *et al.* 2013). These estimates were similar to those of the Indian Ocean island, La Réunion, to which *C. rosa* spread within the last 60 years (De Meyer *et al.* 2008), based on the Baliraine *et al.* (2004) study, but higher than values reported by Virgilio *et al.* (2013). These elevated levels of genetic diversity, even on island populations, indicate that although *C. rosa* has not colonized many locations outside of its native range, when they have successfully established, populations have reached high effective population sizes. This may be at least partly due to their polyphagous nature, having a wide range of hosts available year-round, and in the tropics, continuously favourable weather conditions for development to buffer populations against potential bottleneck events. Congruently, results show evidence of population expansion for the pooled populations. This reservoir of high genetic diversity provides *C. rosa* with elevated levels of evolutionary potential which perhaps goes hand-in-hand with higher levels of phenotypic plasticity, both of which can be considered an adaptive evolutionary advantage in the colonization of novel habitats (Lavergne & Molofsky 2007; Novak 2007; but see discussion in Nyamukondiwa *et al.* 2010 in the context of thermal tolerance plasticity in *Ceratitis*).

Molecular and morphological results clearly indicate a single genotypic cluster and morphotype (R2) present in all sampled locations in South Africa. Virgilio *et al.* (2013) included six South African locations and showed that there was clear genetic structuring between the two *C. rosa* morphotypes (R1 and R2) found sympatrically at two of these locations [Nelspruit (Mpumalanga) and Marblehall (Limpopo) (based on their GPS coordinates)]. Further, they showed that the country seemed to lack pure populations of the R1 morphotype. Based on morphological inspection of the legs of *C. rosa* males, it is clear that both morphotypes (albeit R1 at low abundance) occur sympatrically in locations in the north of South Africa, especially in the Mpumalanga and Limpopo provinces (Tzaneen, Rustenburg, Komatiepoort,

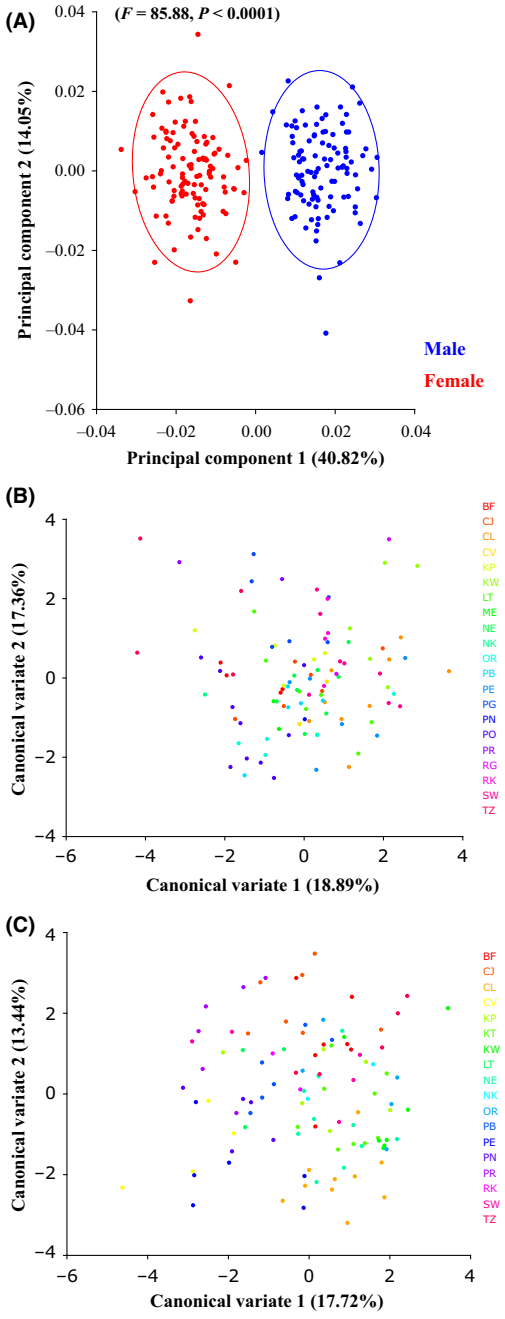

Fig. 4 (A) Principal components analysis (PCA) of wing shape data for *Ceratitis rosa* individuals. Data points coloured according to sex. Ellipses represent 95% equal frequencies. Canonical Variate analysis (CVA) of wing shape data for *C. rosa* individuals by location (ID as in Table 1) for (B) females and (C) males.

Nelspruit, Marblehall) (M. de Meyer, personal communication). These characters on male mid tibia are not influenced by the DNA extraction process and the two morphotypes are easily distinguishable. We therefore expected to find only the one morphotype (R2) over large parts of South Africa, yet flies sampled from the northern locations were also unexpectedly from only a single morphotype. Why only individuals of one morphotype were detected at locations where two morphotypes were found previously is unclear. Possible explanations for this phenomenon may be that the type of host plant in which development takes place may have been absent in the developmental period prior to our sampling as well the season in which individuals were sampled, perhaps suggesting that morphotypes are a consequence of seasonal climatic variability (i.e. developmentally plastic for these phenotypes). For *C. rosa*, the month and year of sampling appear to be factors that might be influencing the abundance of morphotypes, but especially the abundance of the R1 morphotype (M. de Meyer, personal communication). Moreover, recent studies have shown that the two morphotypes are likely adapted to different temperature regimes and therefore occur at different abundances across an altitudinal gradient (Mwatawala *et al.* 2015; Tanga *et al.* 2015). Although we sampled extensively throughout South Africa, failure to trap the R1 morphotype might be a consequence of environmental factors (e.g. temperature) and trapping efficiency if the R1 morphotype is present at lower abundances compared to the R2 morphotype at our sampling locations.

The lack of genetic differentiation between locations within the same morphotype (R2) suggests high levels of dispersal across large distances throughout South Africa which, in turn, may affect the successful implementation of area-wide pest management strategies and could be important in establishment in novel environments. Our data support the view that a management plan is required at a much larger spatial scale than is typically employed (farm or valley), and are further supported by the possibility that there may be human-assisted dispersal in *C. rosa*, similar to what was proposed for *Ceratitis capitata* in South Africa previously (Karsten *et al.* 2013). The need for improved area-wide pest management of fruit flies in South Africa, especially new invasive fruit flies, has also been highlighted in previous studies assessing current management methods (Manrakhan & Addison 2014) and detailing the importance of alternate host plants for the spread of fruit flies (De Villiers *et al.* 2013b). This further suggests that border control may be a critical aspect of fruit fly management and in the prevention of future invasions into currently pest-free regions.

Although we initially set out to ascertain whether wing shape in *C. rosa* can be used to resolve the distribution of the two morphotypes (R1 and R2) in South Africa, geometric morphometrics can also be used to investigate population or ecotype structure. For example, Bouyer *et al.* (2007) showed significant shape differences between different populations of *Glossina palpalis*

gambiensis sampled from Burkina Faso. In *G. palpalis* the differences between the sampling localities based on wing shape were more prominent than the differences found from microsatellites. Based on our genetic results as well as morphological inspection of *C. rosa* individuals, South Africa appears to largely have only one morphotype (R2), with no intraspecific differentiation between sampling localities. Results from the geometric morphometric analyses further supported these findings, although sample sizes for the wing shape analyses were limited in some locations. We also found that *C. rosa* is sexually dimorphic based on wing shape, which is not unexpected. Sexual dimorphism has mainly been investigated as differences in body size (Fairbairn *et al.* 2007). However, in nine species in the *Drosophila melanogaster* subgroup sexual dimorphism based on wing shape was investigated and all nine species exhibited sexual dimorphism (Gidaszewski *et al.* 2009). Sexual dimorphism in tephritids may be linked to courtship behaviour as males use wing movements (among other stimuli) to attract the attention of females (Wicker-Thomas 2007) or sex-related differences in dispersal or flight ability (Sivinski & Dodson 1992; Esterhuizen *et al.* 2014). After separation of males and females, however, there was still no differentiation between sampling localities based on wing shape.

There is a rapidly increasing body of literature (Appendix S7, Supporting information) for species in the Tephritidae showing that these flies are not only widespread but disperse far more than previously thought. The movement of pest fruit flies occurs not only by natural dispersal but is likely also aided by human assistance, for example, through trade or private movement of agricultural produce. This therefore represents a significant management concern, especially in the light of biological invasions, although we are unable to distinguish the relative importance of each of these processes in our data.

Conclusions and future directions

Here, we provide novel evidence that South African populations of *C. rosa* are currently largely representative of only one morphotype (R2) without any significant population differentiation based on both microsatellite data and geometric morphometric data. These results suggest there are high levels of gene flow throughout South Africa. Although these flies can move considerable distances by natural dispersal (several kilometres per generation), aided by the broad distribution of multiple suitable host plants, the possibility that many of these events have been human-mediated is also likely. These high levels of dispersal play an important role in the successful implementation of an IPM program as well as prevention of new invasions, suggesting that population control measures should be implemented over much larger areas than is currently the case. The two morphotypes (R1 and R2) have also been shown to differ in their thermal performance and thermal developmental requirements. The lack of the R1 morphotype may therefore be due to differences in abundance at different temperature regimes. Future work could investigate the possibility of these traits being sex-linked (Nylin *et al.* 1994; Ellers & Boggs 2002; Qvarnström & Bailey 2008) or under frequency-dependent selection (Mallet & Joron 1999; Borer *et al.* 2010).

Acknowledgements

This research was funded by THRIP from the National Research Foundation (NRF) of South Africa [(UID) 71909], South African Apple and Pear Producers' Association (SAAPPA), South African Stone Fruit Producers' Association (SASPA), as well as South African Table Grape Industry (SATGI). The Grantholder acknowledges that opinions, findings and conclusions or recommendations expressed in any publication generated by NRF supported research are that of the author(s), and that the NRF accepts no liability whatsoever in this regard. Genotyping was performed at the Central Analytical Facility, Stellenbosch University. Our thanks to several farmers, Marelize de Villiers and Francois Bekker for technical support and samples. Chris Klingenberg and Hugo Benítez provided valuable input on the geometric morphometrics analyses. We are grateful for constructive comments on this work by the anonymous referees, Sean Schoville, Marc de Meyer and Massimiliano Virgilio.

References

Aketarawong N, Bonizzoni M, Thanaphum S *et al.* (2007) Inferences on the population structure and colonization process of the invasive oriental fruit fly, *Bactrocera dorsalis* (Hendel). *Molecular Ecology*, **16**, 3522–3532.

Aketarawong N, Guglielmino CR, Karam N *et al.* (2014) The oriental fruitfly *Bactrocera dorsalis s.s.* in East Asia: disentangling the different forces promoting the invasion and shaping the genetic make-up of populations. *Genetica*, **142**, 201–213.

Baliraine FN, Bonizzoni M, Guglielmino CR *et al.* (2004) Population genetics of the potentially invasive African fruit fly species, *Ceratitis rosa* and *Ceratitis fasciventris* (Diptera: Tephritidae). *Molecular Ecology*, **13**, 683–695.

Barr NB, McPheron BA (2006) Molecular phylogenetics of the genus *Ceratitis* (Diptera: Tephritidae). *Molecular Phylogenetics & Evolution*, **38**, 216–230.

Barr NB, Copeland RS, De Meyer M *et al.* (2006) Molecular diagnostics of economically important *Ceratitis* fruit fly species (Diptera: Tephritidae) in Africa using PCR and RFLP analyses. *Bulletin of Entomological Research*, **96**, 505–521.

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (1996–2004) *GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations*. Laboratoire Génome, Populations,

Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier.

Berthouly-Salazar C, Janse van Rensburg B, Le Roux JJ, Jansen van Vuuren B, Hui C (2012) Spatial sorting drives morphological variation in the invasive bird, *Acridotheris tristis*. *PLoS ONE*, **7**, e38145.

Berthouly-Salazar C, Hui C, Blackburn TM *et al.* (2013) Long-distance dispersal maximizes evolutionary potential during rapid geographic range expansion. *Molecular Ecology*, **22**, 5793–5804.

Bohonak AJ (1999) Dispersal, gene flow, and population structure. *The Quarterly Review of Biology*, **74**, 21–45.

Bomfim ZV, Lima KM, Silva JG, Costa MA, Zucchi RA (2011) A morphometric and molecular study of *Anastrepha pickeli* Lima (Diptera: Tephritidae). *Neotropical Entomology*, **40**, 587–594.

Bonizzoni M, Zheng L, Guglielmino CR *et al.* (2001) Microsatellite analysis of medfly bioinfestations in California. *Molecular Ecology*, **10**, 2515–2524.

Bonizzoni M, Guglielmino CR, Smallridge CJ, Gomulski M, Malacrida AR, Gasperi G (2004) On the origins of medfly invasion and expansion in Australia. *Molecular Ecology*, **13**, 3845–3855.

Bookstein FL (1991) *Morphometric Tools for Landmark Data: Geometry and Biology*. Cambridge University Press, New York, New York.

Borer M, Van Noort T, Rahier M, Naisbit RE (2010) Positive frequency-dependent selection on warning color in alpine leaf beetles. *Evolution*, **64**, 3629–3633.

Bouyer J, Ravel S, Dujardin JP *et al.* (2007) Population structuring of *Glossina palpalis gambiensis* (Diptera: Glossinidae) according to landscape fragmentation in the Mouhoun river, Burkina Faso. *Journal of Medical Entomology*, **44**, 788–795.

Bouyer J, Balenghien T, Ravel S *et al.* (2009) Population sizes and dispersal pattern of tsetse flies: rolling on the river? *Molecular Ecology*, **18**, 2787–2797.

Chapuis MP, Estoup A (2007) Microsatellite null alleles and estimation of population differentiation. *Molecular Biology & Evolution*, **24**, 621–631.

Chapuis MP, Popple JAM, Berthier K *et al.* (2011) Challenges to assessing connectivity between massive populations of the Australian plague locust. *Proceedings of the Royal Society of London B: Biological Sciences*, **278**, 3152–3160.

Copeland RS, Wharton RA, Luke Q *et al.* (2006) Geographic distribution, host fruit, and parasitoids of African fruit fly pests *Ceratitis anonae*, *Ceratitis cosyra*, *Ceratitis fasciventris*, and *Ceratitis rosa* (Diptera: Tephritidae) in Kenya. *Annals of the Entomological Society of America*, **99**, 261–278.

De Meyer M, Copeland RS, Wharton RA, McPheron BA (2002) On the geographic origin of the Medfly *Ceratitis capitata* (Weidemann) (Diptera: Tephritidae). *Proceedings of the 6th International Fruit Fly Symposium*, pp. 45–53.

De Meyer M, Robertson MP, Peterson AT, Mansell MW (2008) Ecological niches and potential geographical distributions of Mediterranean fruit fly (*Ceratitis capitata*) and Natal fruit fly (*Ceratitis rosa*). *Journal of Biogeography*, **35**, 270–281.

De Meyer M, Delatte H, Ekesi S *et al.* (2015) An integrative approach to unravel the *Ceratitis* FAR (Diptera, Tephritidae) cryptic species complex: a review. *ZooKeys*, **540**, 405–427.

De Villiers M, Hattingh V, Kriticos DJ (2013a) Combining field phenological observations with distribution data to model the potential distribution of the fruit fly *Ceratitis rosa* Karsch (Diptera: Tephritidae). *Bulletin of Entomological Research*, **103**, 60–73.

De Villiers M, Manrakhan A, Addison P, Hattingh V (2013b) The distribution, relative abundance, and seasonal phenology of *Ceratitis capitata*, *Ceratitis rosa*, and *Ceratitis cosyra* (Diptera: Tephritidae) in South Africa. *Environmental Entomology*, **42**, 831–840.

Delatte H, Virgilio M, Simiand C, Quilici S, De Meyer M (2013) Isolation and characterisation of sixteen microsatellite markers cross-amplifying in a complex of three African agricultural pests (*Ceratitis rosa*, *C. anonae* and *C. fasciventris*, Diptera: Tephritidae). *Conservation Genetics Resources*, **5**, 31–34.

Dempster AP, Laird NM, Rubin DB (1977) Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society*, **39**, 1–38.

Dieringer D, Schlötterer C (2003) Microsatellite analyser (MSA): a platform independent analysis tool for large microsatellite data sets. *Molecular Ecology Notes*, **3**, 167–169.

Do C, Waples RS, Peel D, Macbeth GM, Tillett BJ, Ovenden JR (2014) NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. *Molecular Ecology Resources*, **14**, 209–214.

Douglas LJ, Haymer DS (2001) Ribosomal ITS1 polymorphisms in *Ceratitis capitata* and *Ceratitis rosa* (Diptera: Tephritidae). *Annals of the Entomological Society of America*, **94**, 726–731.

Duncan RP, Blackburn TM, Rossinelli S, Bacher S (2014) Quantifying invasion risk: the relationship between establishment probability and founding population size. *Methods in Ecology & Evolution*, **5**, 1255–1263.

Duyck PF, Quilici S (2002) Survival and development of different life stages of three *Ceratitis* spp. (Diptera: Tephritidae) reared at five constant temperatures. *Bulletin of Entomological Research*, **92**, 461–469.

Duyck PF, David P, Quilici S (2006) Climatic niche partitioning following successive invasions by fruit flies in La Réunion. *Journal of Animal Ecology*, **75**, 518–526.

Earl DA, Von Holdt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, **4**, 359–361.

Ellers J, Boggs CL (2002) The evolution of wing color in *Colias* butterflies: heritability, sex linkage, and population divergence. *Evolution*, **56**, 836–840.

Esterhuizen N, Clusella-Trullas S, van Daalen CE, Schoombie RE, Boardman L, Terblanche JS (2014) Effects of within-generation thermal history on flight performance of *Ceratitis capitata*: colder is better. *Journal of Experimental Biology*, **217**, 3545–3556.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611–2620.

Excoffier L, Lischer HEL (2010) Arlequin suite version 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564–567.

Fairbairn DJ, Blanckenhorn W, Szekely T (2007) *Sex, Size and Gender Roles: Evolutionary Studies of Sexual Size Dimorphism*. Oxford University Press, New York.

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, **164**, 1567–1587.

Forester BR, Jones MR, Joost S, Landguth EL, Lasky JR (2016) Detecting spatial genetic signatures of local adaptation in heterogeneous landscapes. *Molecular Ecology*, **25**, 104–120.

Francoy TM, Silva RAO, Nunes-Silva P, Menezes C, Imperatriz-Fonseca VL (2009) Gender identification of five genera of stingless bees (Apidae, Meliponini) based on wing morphology. *Genetics & Molecular Research*, **8**, 207–214.

Francuski L, Djurakic M, Ståhls G, Milankov V (2014) Landscape genetics and wing morphometrics show a lack of structuring across island and coastal populations of the drone fly in the Mediterranean. *Journal of Zoology*, **292**, 156–169.

Gidaszewski NA, Baylac M, Klingenberg CP (2009) Evolution of sexual dimorphism of wing shape in the *Drosophila melanogaster* subgroup. *BMC Evolutionary Biology*, **9**, 110–120.

Gilchrist AS, Crisafulli DCA (2006) Using variation in wing shape to distinguish between wild and mass-reared individuals of Queensland fruit fly, *Bactrocera tryoni*. *Entomologia Experimentalis et Applicata*, **119**, 175–178.

Gilchrist AS, Meats AW (2009) The genetic structure of populations of an invading pest fruit fly, *Bactrocera tryoni*, at the species climatic range limit. *Heredity*, **105**, 165–172.

Hill MP, Bertelsmeier C, Clusella-Trullas S, Garnas J, Robertson MP, Terblanche JS (2016) Predicted decrease in global climate suitability masks regional complexity of invasive fruit fly species response to climate change. *Biological Invasions*, **18**, 1105–1119.

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, **23**, 1801–1806.

Karsten M, Jansen van Vuuren B, Barnaud A, Terblanche JS (2013) Population genetics of *Ceratitis capitata* in South Africa: implications for dispersal and pest management. *PLoS ONE*, **8**, e54281.

Karsten M, Jansen van Vuuren B, Addison P, Terblanche JS (2015) Deconstructing intercontinental invasion pathway hypotheses of the Mediterranean fruit fly (*Ceratitis capitata*) using a Bayesian inference approach: are port interceptions and quarantine protocols successfully preventing new invasions? *Diversity and Distributions*, **21**, 813–825.

Kawecki TJ, Ebert D (2004) Conceptual issues in local adaptation. *Ecology Letters*, **7**, 1225–1241.

Klingenberg CP (2011) MorphoJ: an integrated software package for geometric morphometrics. *Molecular Ecology Resources*, **11**, 353–357.

Krosch MN, Schutze MK, Armstrong KF *et al.* (2013) Piecing together an integrative taxonomic puzzle: microsatellite, wing shape and aedeagus length analyses of *Bactrocera dorsalis s.l.* (Diptera: Tephritidae) find no evidence of multiple lineages in a proposed contact zone along the Thai/Malay Peninsula. *Systematic Entomology*, **38**, 2–13.

Lavergne S, Molofsky J (2007) Increased genetic variation and evolutionary potential drive the success of an invasive grass. *Proceedings of the National Academy of Sciences*, **104**, 3883–3888.

Leblois R, Pudlo P, Néron J *et al.* (2014) Maximum likelihood inference of population size contractions from microsatellite data. *Molecular Biology and Evolution*, **31**, 2805–2823.

Lenormand T (2002) Gene flow and the limits to natural selection. *Trends in Ecology & Evolution*, **17**, 183–189.

Lockwood JL, Cassey P, Blackburn TM (2009) The more you introduce the more you get: the role of colonization pressure and propagule pressure in invasion ecology. *Diversity & Distributions*, **15**, 904–910.

Mallet J, Joron M (1999) Evolution of diversity in warning coloration and mimicry: polymorphisms, shifting balance, and speciation. *Annual Review of Ecology & Systematics*, **30**, 201–233.

Manrakhan A, Addison P (2014) Assessment of fruit fly (Diptera: Tephritidae) management practices in deciduous fruit growing areas in South Africa. *Pest Management Science*, **70**, 651–660.

Marsteller S, Adams DC, Collyer ML, Condon M (2009) Six cryptic species on a single species of host plant: morphometric evidence for possible reproductive character displacement. *Ecological Entomology*, **34**, 66–73.

Meats A, Smallridge CJ (2007) Short- and long-range dispersal of medfly, *Ceratitis capitata* (Dipt., Tephritidae), and its invasive potential. *Journal of Applied Entomology*, **131**, 518–523.

Mwatawala M, Virgilio M, Joseph J, De Meyer M (2015) Niche partitioning among two *Ceratitis rosa* morphotypes and other *Ceratitis* pest species (Diptera, Tephritidae) along an altitudinal transect in Central Tanzania. *ZooKeys*, **540**, 429–442.

Nardi F, Carapelli A, Dallai R, Roderick GK, Frati F (2005) Population structure and colonization history of the olive fly, *Bactrocera oleae* (Diptera, Tephritidae). *Molecular Ecology*, **14**, 2729–2738.

Novak SJ (2007) The role of evolution in the invasion process. *Proceedings of the National Academy of Sciences*, **104**, 3671–3672.

Nyamukondiwa C, Kleynhans E, Terblanche JS (2010) Phenotypic plasticity of thermal tolerance contributes to the invasion potential of Mediterranean fruit flies (*Ceratitis capitata*). *Ecological Entomology*, **35**, 565–575.

Nylin S, Wickman P-O, Wiklund C (1994) Genetics of development time in a butterfly: redictions from optimality and a test by subspecies crossings. *Proceedings of the Royal Society of London B: Biological Sciences*, **257**, 215–219.

Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, **6**, 288–295.

Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research —an update. *Bioinformatics*, **28**, 2537–2539.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

Qvarnström A, Bailey RI (2008) Speciation through evolution of sex-linked genes. *Heredity*, **102**, 4–15.

Raymond M, Rousset F (1995) An exact test for population differentiation. *Evolution*, **49**, 1280–1283.

Richter-Boix AL, Teplitsky C, Rogell B, Laurila A (2010) Local selection modifies phenotypic divergence among *Rana temporaria* populations in the presence of gene flow. *Molecular Ecology*, **19**, 716–731.

Rohlf FJ (1999) Shape statistics: procrustes superimpositions and tangent spaces. *Journal of Classification*, **16**, 197–223.

Rohlf FJ (2005) *tpsDig, Digitize Landmarks and Outlines, Version 2.05*. Department of Ecology and Evolution, State University of New York at Stony Brook, Stony Brook, New York.

Rohlf FJ, Marcus LF (1993) A revolution in morphometrics. *Trends in Ecology & Evolution*, **8**, 129–132.

Rosenberg NA (2004) DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes*, **4**, 137–138.

Rousset F (2008) GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources*, **8**, 103–106.

Schutze MK, Jessup A, Clarke AR (2012a) Wing shape as a potential discriminator of morphologically similar pest taxa within the *Bactrocera dorsalis* species complex (Diptera: Tephritidae). *Bulletin of Entomological Research*, **102**, 103–111.

Schutze MK, Krosch MN, Armstrong KF *et al.* (2012b) Population structure of *Bactrocera dorsalis ss*, *B. papayae* and *B. philippinensis* (Diptera: Tephritidae) in southeast Asia: evidence for a single species hypothesis using mitochondrial DNA and wing-shape data. *BMC Evolutionary Biology*, **12**, 130.

Schutze MK, Mahmood K, Pavasovic A *et al.* (2015a) One and the same: integrative taxonomic evidence that *Bactrocera invadens* (Diptera: Tephritidae) is the same species as the Oriental fruit fly *Bactrocera dorsalis*. *Systematic Entomology*, **40**, 472–486.

Schutze MK, Aketarawong N, Amornsak W *et al.* (2015b) Synonymization of key pest species within the *Bactrocera dorsalis* species complex (Diptera: Tephritidae): taxonomic changes based on a review of 20 years of integrative morphological, molecular, cytogenetic, behavioural and chemoecological data. *Systematic Entomology*, **40**, 456–471.

Sexton JP, Hangartner SB, Hoffmann AA (2014) Genetic isolation by environment or distance: which pattern of gene flow is most common? *Evolution*, **68**, 1–15.

Sivinski JM, Dodson G (1992) Sexual dimorphism in *Anastrepha suspensa* (Loew) and other tephritid fruit flies (Diptera: Tephritidae): possible roles of developmental rate, fecundity, and dispersal. *Journal of Insect Behavior*, **5**, 491–506.

Storey JD (2002) A direct approach to false discovery rates. *Journal of the Royal Statistical Society, Series B*, **64**, 479–498.

Szűcs M, Melbourne BA, Tuff T, Hufbauer RA (2014) The roles of demography and genetics in the early stages of colonization. *Proceedings of the Royal Society of London B: Biological Sciences*, **281**, 20141073.

Tanga CM, Manrakhan A, Daneel JH, Mohamed SA, Fathiya K, Ekesi S (2015) Comparative analysis of development and survival of two Natal fruit fly *Ceratitis rosa* Karsch (Diptera, Tephritidae) populations from Kenya and South Africa. *ZooKeys*, **540**, 467–487.

Vaníčková L, Břízová R, Pompeiano A, Ekesi S, De Meyer M (2015) Cuticular hydrocarbons corroborate the distinction between lowland and highland Natal fruit fly (Tephritidae, *Ceratitis rosa*) populations. *ZooKeys*, **540**, 507–524.

Virgilio M, Backeljau T, Barr N, De Meyer M (2008) Molecular evaluation of nominal species in the *Ceratitis fasciventris*, *C. anonae*, *C. rosa* complex (Diptera: Tephritidae). *Molecular Phylogenetics & Evolution*, **48**, 270–280.

Virgilio M, Delatte H, Backeljau T, De Meyer M (2010) Macrogeographic population structuring in the cosmopolitan agricultural pest *Bactrocera cucurbitae* (Diptera: Tephritidae). *Molecular Ecology*, **19**, 2713–2724.

Virgilio M, Delatte H, Quilici S, Backeljau T, De Meyer M (2013) Cryptic diversity and gene flow among three African agricultural pests: *Ceratitis rosa*, *Ceratitis fasciventris* and *Ceratitis anonae* (Diptera, Tephritidae). *Molecular Ecology*, **22**, 2526–2539.

Von Wettberg EJ, Remington DL, Schmitt J (2008) Partitioning adaptive differentiation across a patchy landscape: shade avoidance traits in *Impatiens capensis*. *Evolution*, **62**, 654–667.

Waples RS, Do CHI (2008) LDNE: a program for estimating effective population size from data on linkage disequilibrium. *Molecular Ecology Resources*, **8**, 753–756.

White IM, Elson-Harris MM (1994) *Fruit Flies of Economics Significance: Their Identification and Bionomics*. ACIAR, CAB International, Wallingford, Oxon.

White IM, De Meyer M, Stonehouse J (2001) A review of the native and introduced fruit flies (Diptera: Tephritidae) in the Indian Ocean Islands of Mauritius, Reunion, Rodrigues and Seychelles. *Proceedings of the Indian Ocean Commission Regional Fruit Fly Symposium*, pp. 15–21.

Wicker-Thomas C (2007) Pheromonal communication involved in courtship behavior in Diptera. *Journal of Insect Physiology*, **53**, 1089–1100.

M.K. contributed to the study design, conducted sampling, performed molecular, geometric and statistical analyses and drafted the manuscript. J.S.T. contributed to study design, provided reagents, materials and analysing tools and writing of the manuscript. P.A. contributed to geometric analysis, reagents, materials and analysing tools and the revision of the manuscript. B.J.V.V. contributed to preparation and revision of the manuscript.

Data accessibility

Microsatellite genotypes and wing shape data available on DRYAD: doi: 10.5061/dryad.sn62j.

Supporting information

Additional supporting information may be found in the online version of this article.

Appendix S1 Right wing of *Ceratitis rosa* showing the fourteen type I landmarks (Bookstein 1991) used in the geometric morphometric data set.

Appendix S2 Estimates of null allele frequencies (for each locus in each population, averaged over populations and averaged over loci) of *Ceratitis rosa* calculated based on the EM algorithm implemented in the program FREENA (Chapuis & Estoup 2007).

Appendix S3 Linear correlation of genetic distance (pairwise $F_{ST}/1-F_{ST}$) against geographic distance (ln distance in km) between sampling locations.

Appendix S4 Analysis of molecular variance (AMOVA) for *Ceratitis rosa* in South Africa (a) based on microsatellites calculated in ARLEQUIN v3.5.1.2 (Excoffier & Lischer 2010) and (b) a Procrustes ANOVA of shape variation for the error dataset based on a subset of our *Ceratitis rosa* wing shape measurements.

Appendix S5 Pairwise F_{ST} values calculated from 11 polymorphic microsatellite markers for *Ceratitis rosa* sampled across South Africa.

Appendix S6 Linear correlation of Mahalanobis distance (from the CVA) against geographic distance between sampling locations (in km) (A) in males and (B) in females.

Appendix S7 A selection of population genetic structure studies in the Tephritidae.